

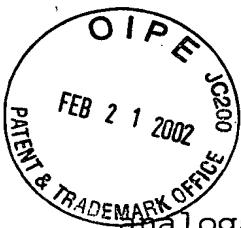


by standard methods (Sambrook et al., Molecular Cloning, CSH Press (1987)).

Any of the formats for in vitro or in vivo recombination described herein can be applied for recursive 5 exon shuffling. A preferred format is to use nonpalindromic restriction sites such as Sfi I placed into the intronic sequences to promote shuffling. Pools of selected clones are digested with Sfi I and religated. The nonpalindromic overhangs promote ordered reassembly of the shuffled exons.

10 These libraries of genes can be expressed and screened for desired properties, then subjected to further recursive rounds of recombination by this process. In some embodiments, the libraries are subjected to mutagenesis before the process is repeated.

15 An example of how the introduction of an intron into a mammalian library format would be used advantageously is as follows. An intron containing a lox (Sauer et al., Proc. Natl. Acad. Sci. (U.S.A.), 85:5166-5170 (1988)) site is arbitrarily introduced between amino acids 92 and 93 in each alpha interferon parental substrate. A library of 10^4 chimeric interferon genes is made for each of the two exons (residues 1-92 and residues 93-167), cloned into a replicating plasmid vector, and introduced into target cells. The number 20 10^4 is arbitrarily chosen for convenience in screening. An exemplary vector for expression in mammalian cells would contain an SV40 origin, with the host cells expressing SV40 large T antigen, so as to allow transient expression of the interferon constructs. The cells are challenged with a cytopathic virus such as vesicular stomatitis virus (VSV) in 25 an interferon protection assay (e.g., Meister et al., J. Gen. Virol. 67:1633-1643, (1986)). Cells surviving due to expression of interferon are recovered, the two libraries of interferon genes are PCR amplified, and recloned into a vector that can be amplified in *E. coli*. The amplified plasmids are 30 then transfected at high multiplicity (e.g. 10 micrograms of plasmid per 10^6 cells) into a cre expressing host that can support replication of that vector. The presence of cre in 35



analogs such as inosine, 7-deaza dGTP (Dierick et al., Nucleic Acids Res. 21:4427-4428 (1993)) or other novel base analogs that improve the above properties.

Novel sequencing formats have been described which use
5 matrix assisted laser desorption ionization time of flight (MALDI-TOF) mass spectroscopy to resolve dideoxy ladders (Smith, Nature Biotechnology 14:1084-1085 (1996)). It is noted in Smith's recent review that fragmentation of the DNA is the singular feature limiting the development of this method as a viable alternative to
10 standard gel electrophoresis for DNA sequencing. Base analogs which stabilize the N-glycosidic bond by modifications of the purine bases to 7-deaza analogs (Kirpekar et al., Rapid Comm. in Mass Spec. 9:525-531 (1995)) or of the 2' hydroxyl (such as 2'-H or 2'-F) "relieve greatly the mass range limitation" of this technique (Smith, 1996). Thus, evolved polymerases that can efficiently incorporate these and other base analogs conferring resistance to fragmentation under MALDI-TOF conditions are
15 valuable innovations.

Other polymerase properties of interest for improvement by RSR are low fidelity thermostable DNA polymerase for more efficient mutagenesis or as a useful correlate for acceptance of base analogs for the purposes described above; higher fidelity polymerase for PCR (Lundberg et al., Gene 108:1-6 (1991)); higher fidelity reverse transcriptase for retroviral gene therapy
20 vehicles to reduce mutation of the therapeutic construct and of the retrovirus; improved PCR of GC rich DNA and PCR with modified bases (S. Turner and F. J. Jenkins, BioTechniques 19(1):48-52 (1995)).

Thus, in some embodiments of the invention, libraries of
30 mutant polymerase genes are screened by direct high throughput screening for improved sequencing properties. The best candidates are then subjected to RSR. Briefly, mutant libraries of candidate polymerases such as Taq polymerase are constructed using standard methods such as PCR mutagenesis (Caldwell et al., PCR Meth. App. 2:28-33 (1992)) and/or cassette mutagenesis (Sambrook et al., Molecular Cloning, CSH Press (1987)). Incorporation of mutations
35 into Taq DNA polymerase such as the active site residue from T7



sequencing, PCR of GC rich DNA (>60% GC), PCR with novel base analogs such as 7-deaza purines, 2' fluoro dNTP's, rNTP's, PCR with inosine, etc.). The amplified genes are pooled, cloned, and subjected to mutagenesis, and the process repeated until an improvement is achieved.

5 C. Evolved Phosphonatase

10
15
20
25
30
35

Alkaline phosphatase is a widely used reporter enzyme for ELISA assays, protein fusion assays, and in a secreted form as a reporter gene for mammalian cells. The chemical lability of p-nitrophenyl phosphate (pNPP) substrates and the existence of cellular phosphatases that cross-react with pNPP is an important limitation on the sensitivity of assays using this reporter gene. A reporter gene with superior signal to noise properties can be developed based on hydrolysis of p-nitrophenyl phosphonates, which are far more stable to base catalyzed hydrolysis than the corresponding phosphates. Additionally, there are far fewer naturally occurring cellular phosphonatases than alkaline phosphatases. Thus a p-nitrophenyl phosphonatase is an attractive replacement for alkaline phosphatase because the background due to chemical and enzymatic hydrolysis is much lower. This will allow one to make ELISA's more sensitive for detecting very small concentrations of antigen.

Chen et al. (J. Mol. Biol. 234:165-178 (1993)) have shown that a *Staph. aureus* beta-lactamase can hydrolyze p-nitrophenyl phosphonate esters with single turnover kinetics. The active site Ser70 (the active site nucleophile for beta lactam hydrolysis) forms a covalent intermediate with the substrate. This is analogous to the first step in hydrolysis of beta lactams, and this enzyme can be evolved by RSR to hydrolyze phosphonates by a mechanism analogous to beta lactam hydrolysis. Metcalf and Wanner have described a cryptic phosphonate utilizing operon (phn) in *E. coli*, and have constructed strains bearing deletions of the phn operon (J. Bact. 175:3430-3442 (1993)). This paper discloses selections for growth of *E. coli* on phosphate free minimal media where the phosphorous is derived from hydrolysis of alkyl phosphonates by genes in the phn operon. Thus, one could select